Metagenome Bins
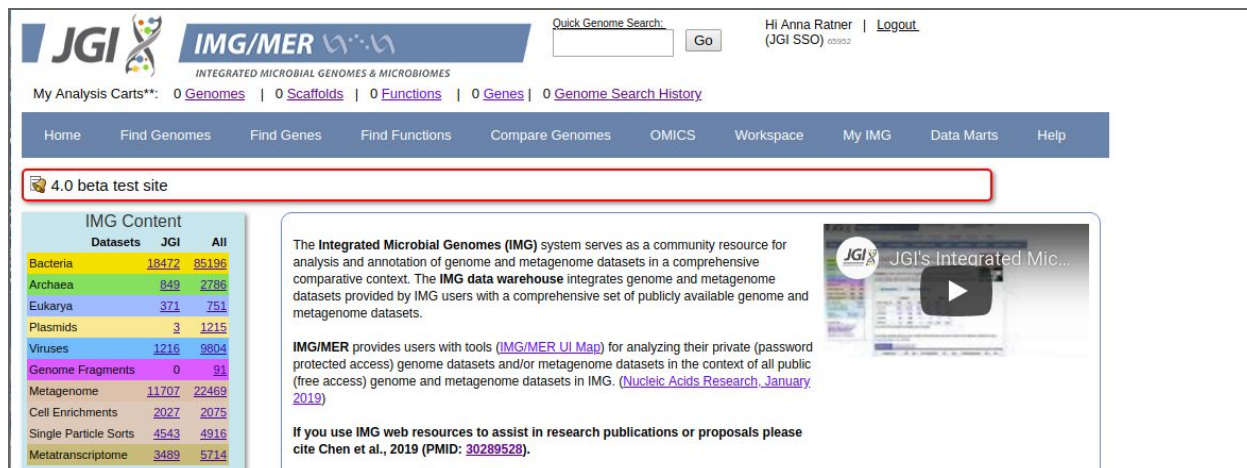
Introduction
Metagenome bins can be accessed in IMG in a variety of ways:
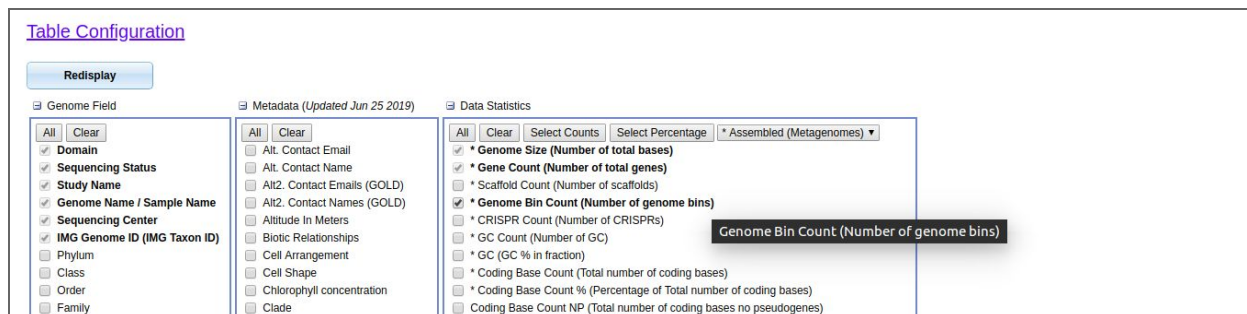
1. Access via Table of Metagenomes
One way to access bins is directly from a table of metagenomes e.g. by clicking on the count of JGI Metagenomes on the IMG home page under IMG Content.

When the table of metagenomes comes up, the user needs to select "**Genome Bin Count**" from the Table Configuration option below the table and click "Redisplay".

The table will then show the <u>count of bins</u> for each metagenome.

The bin count links to <u>bin detail page</u> with links to scaffolds and genes for each metagenome bin.

Metagenome Bins

Genome: <u>Hot spring thermophilic microbial communities from Obsidian Pool, Yellowstone National Park, USA - OP-RAMG-01 (SPAdes)</u>

[Save Selected Bins as Scaffold Sets] [Select All] [Clear All]

Filter column: [Bin ID ▼] Filter [text ▼] : [   ] [Apply] ⓘ

[Export] Page 1 of 3 << first < prev **1** | 2 | 3 | next > last >> [100 ▼]

[Column Selector] [Select Page] [Deselect Page]

Select	Bin ID ▲	Bin Quality	Bin Lineage	Bin Completeness	Bin Contamination	Total Number of Bases	5s rRNA	16s rRNA	23s rRNA	tRNA Genes	Gene Count	Scaffold Count
☐	3300027863_100	MQ	Bacteria	75.59	0.54	2333895	0	0	0	33	2294	361
☐	3300027863_101	MQ	Bacteria	89.45	0.18	2321820	1	1	0	45	2119	117
☐	3300027863_102	HQ	Bacteria	92.74	4.95	2286351	1	1	1	44	2208	112
☐	3300027863_103	MQ	Bacteria	83.62	3.45	2275490	1	0	1	38	2269	288
☐	3300027863_104	MQ	Bacteria	50.09	7.14	2263522	0	0	0	25	2333	456
☐	3300027863_105	HQ	Bacteria	90.74	0	2266204	1	2	1	44	2164	226
☐	3300027863_107	MQ	Bacteria	97.63	0	2249348	2	0	0	34	2183	172
☐	3300027863_108	MQ	Bacteria ; Actinobacteria	95.44	1.28	2235363	1	0	1	45	1922	55
☐	3300027863_11	HQ	Bacteria	92.59	1.75	5171146	3	1	2	64	4599	211
☐	3300027863_110	MQ	Bacteria	60.34	0	2240746	0	1	1	38	2197	156
☐	3300027863_112	MQ	Bacteria	89.41	5.91	2213099	1	1	0	42	2183	202
☐	3300027863_113	MQ	Bacteria	78.85	7.79	2193531	0	0	1	32	2047	316
			Bacteria ; Proteobacteria ;									

2. Access via Metagenome Detail Page

If the user is interested in a particular metagenome which has bins, the bins can be accessed directly from the metagenome details page.

Microbiome Details (Assembled Data)

[Add to Genome Cart] [🖼 Browse Genome] [ATC ATC BLAST Genome] [📄 Download Data]

About Genome

- <u>Overview</u>
- <u>Statistics</u>
- <u>Genes</u>

Overview

Study Name (Proposal Name)	Freshwater sediment methanotrophic microbial communities from Lake Washington under simulated oxygen tension
Sample Name	Freshwater sediment methanotrophic microbial communities from Lake Washington under simulated oxygen tension - Sediment Metagenome 20_HOW5 (SPAdes)
Taxon Object ID	3300024970
IMG Submission ID	183866
GOLD ID in IMG Database	
GOLD Analysis Project Id	Ga0207967
GOLD Analysis Project Type	Metagenome Analysis
Submission Type	Reanalysis
JGI Analysis Product Name	Metagenome Minimal Draft
JGI Analysis Project Type	Metagenome Analysis
Sequencing Status	Draft
Sequencing Center	DOE Joint Genome Institute (JGI)
IMG Release/Pipeline Version	IMG Annotation Pipeline v.4.16.2

CRISPR Count	125	-
Genes		
RNA genes	1115	0.69%
rRNA genes	189	0.12%
5S rRNA	48	0.03%
16S rRNA	55	0.03%
18S rRNA	6	0.00%
23S rRNA	76	0.05%
28S rRNA	4	0.00%
tRNA genes	926	0.57%
Protein coding genes	161581	99.31%
with Product Name	113715	69.89%
with COG	100026	61.48%
with Pfam	97812	60.12%
with KO	80968	49.77%
with Enzyme	43842	26.95%
with MetaCyc	26180	16.09%
with KEGG	46017	28.28%
COG Clusters	3350	72.34%
Pfam Clusters	4591	28.16%
Metagenome Bins	4	

3. Access via Overview of Bins

Metagenome Bins can be accessed in a more efficient manner from the "Find Genomes" menu. They can be viewed in bulk either by Taxonomy or by Ecosystem. Both views present an overview of *all* bins in IMG in a **partition viewer**.

The <u>horizontal partition layout</u> is a way to represent a tree hierarchy where each node is drawn as a rectangle and its children are drawn as smaller adjacent rectangle branches. The height of all the child rectangles matches the height of the parent rectangle. The placement and height of each child node corresponds to the number of leaf nodes in its branch of the tree.

To navigate within the partition, the user can:

1. Click on a parent rectangle of interest to expand that category. The legend will then display a list of all the child nodes for that category, even those not visible in the viewer itself (this happens if the leaf count for the category is relatively low).
2. User can click on an item in the <u>legend</u> to expand that category.
3. User can click on any item in the <u>breadcrumbs</u> (displayed when hovering with the mouse over the partition) to see all items for that category in a bin details table.
4. User can click to the <u>left of a category</u> in the viewer to return to the parent category.

Bins by Taxonomy

> Home > Find Genomes > Bins by Taxonomy

Metagenome Bins can also be **viewed grouped in a table by a category** of choice. The user can select from a <u>dropdown list</u> of categories. This table displays the count of bins for each item in that category "hierarchy". The count links to <u>bin details</u> for that category hierarchy. As noted in the tool, duplicate category names may appear in the table. These would have a distinct hierarchy for the corresponding bins.

Metagenome Bins by Phylogenetic Category

View by Phylogenetic Category:
[Phylum ▾]

> **hint:** *Click on the count to view the bin details for selected category.*
> *To group bins by a different category, choose from the dropdown above.*
> *Select 'Reset' to go back to view all categories in the partition viewer.*
> <u>*Note:*</u> *Duplicate category names may appear. These would have a distinct hierarchy for the corresponding bins.*

Show [10 ▾] entries   Search: [therm   ✕]

Phylum ▲	Bin Count ⇕
Coprothermobacterota	5
Deinococcus-Thermus	211
Rhodothermaeota	4
Thermodesulfobacteria	25
Thermotogae	188

Showing 1 to 5 of 5 entries (filtered from 66 total entries)   Previous [1] Next

Bins by Ecosystem

> Home > Find Genomes > Bins by Ecosystem

Metagenome Bins by Ecosystem

Total Metagenomes with Bins: 7820
Total Bins: 67592

View by Ecosystem Category:
[--- Select a Category --- ▾]

> **hint:** <u>*Partition:*</u> *click on a category in the legend to expand hierarchy under that category (the legend appears when a category is selected). Click on any category in the partition display to expand that category. Click to the left of that category to return to the parent category. Click on a breadcrumb category (displayed when hovering with the mouse over the partition) to display the list of items for the hierarchy up to that category.*
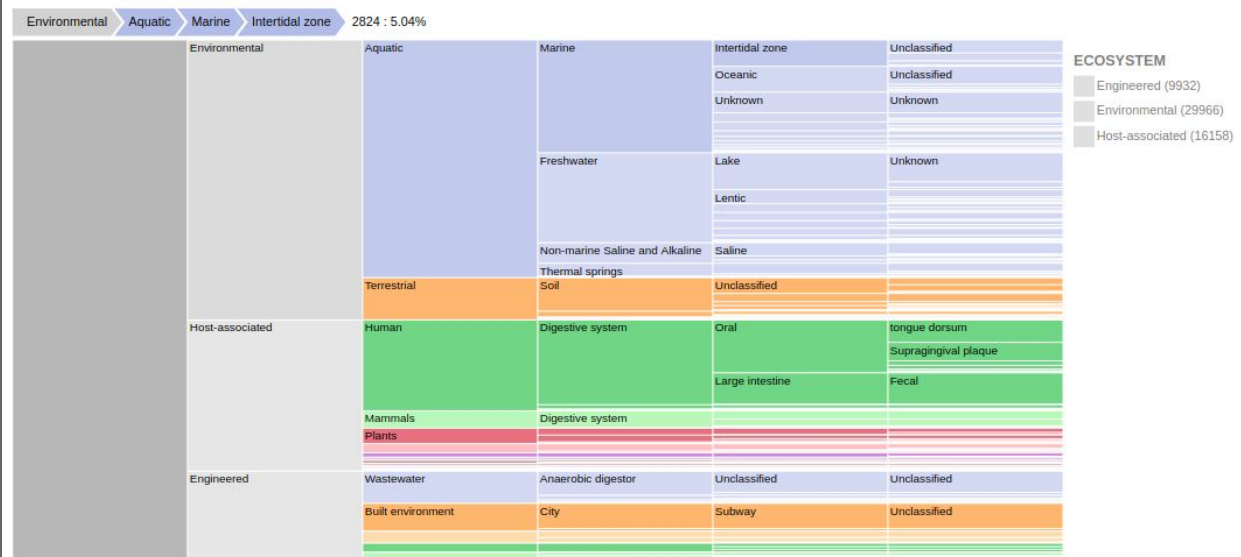
Environmental ⟩ Aquatic ⟩ Marine ⟩ Intertidal zone   2824 : 5.04%

ECOSYSTEM
- Engineered (9932)
- Environmental (29966)
- Host-associated (16158)

Metagenome Bins by Ecosystem

Total Metagenomes with Bins: 7820
Total Bins: 67592

View by Ecosystem Category:

--- Select a Category --- ▼

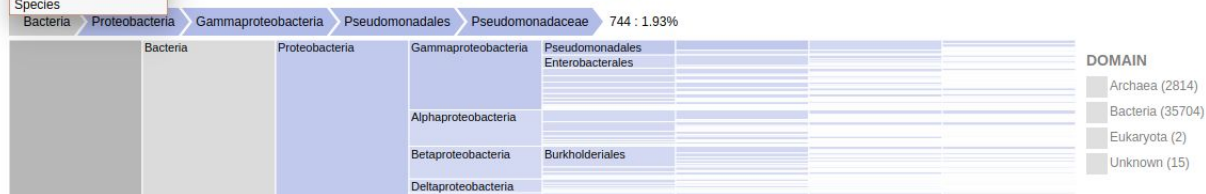| --- Select a Category --- |
| Reset |
| Ecosystem |
| Ecosystem Category |
| Ecosystem Type |
| Ecosystem Subtype |
| Specific Ecosystem |

...on a category in the legend to expand hierarchy under that category (the ... when a category is selected). Click on any category in the partition display to ...k to the left of that category to return to the parent category. Click on a ...played when hovering with the mouse over the partition) to display the list of ... to that category.

| Environmental | Aquatic | 25432 : 45.4% |

	Environmental	Aquatic	Marine	Intertidal zone	Unclassified
				Oceanic	Unclassified
				Unknown	Unknown
			Freshwater	Lake	Unknown

ECOSYSTEM
- Engineered (9932)
- Environmental (29966)
- Host-associated (16158)

Metagenome Bins by Ecosystem Category

View by Ecosystem Category:

Ecosystem Category ▼

hint: *Click on the count to view the bin details for selected category.*
To group bins by a different category, choose from the dropdown above.
Select 'Reset' to go back to view all categories in the partition viewer.
Note: Duplicate category names may appear. These would have a distinct hierarchy for the corresponding bins.

Show 10 ▼ entries   Search: [   ]

Ecosystem_category ▲	Bin Count
Air	23
Algae	323
Animal	451
Annelida	142
Aquatic	25432
Arthropoda	984
Bioreactor	1281
Bioremediation	137
Biotransformation	386
Birds	107

Showing 1 to 10 of 29 entries   Previous 1 2 3 Next

Bins Details Table

The details table for any group of bins lists the bin IDs and links to associated genome, genes, and scaffolds:

Metagenome Bins by Ecosystem

Ecosystem: Host-associated
Ecosystem Category: Annelida

[Save Selected Bins as Scaffold Sets]   [Select All]   [Clear All]

Filter column: Bin ID ▼ Filter text ▼ : [   ] [Apply] ❓

[Export] Page 1 of 2 << first < prev 1 | 2 | next > last >> 100 ▼

[Column Selector] [Select Page] [Deselect Page]

Select	Bin ID ▲	Genome Name	Ecosystem	Ecosystem Type	Ecosystem Subtype	Specific Ecosystem	Bin Quality	Bin Lineage	Bin Completeness	Bin Contamination	Total Number of Bases	5s rRNA	16s rRNA	23s rRNA	tRNA Genes	Gene Count	Scaffold Count
☐	2004178001_1	Olavius algarvensis microbial communities from Capo di Sant'Andrea bay, Elba, Italy - Delta1	Host-associated	Annelida	Integument	Unclassified	MQ	Bacteria ; Proteobacteria ; Deltaproteobacteria ; unclassified ; unclassified ; unclassified	87.63	3.23	12744768	1	1	1	45	11503	198
☐	3300003129_2	Olavius algarvensis microbial community from off the coast of Capo di Sant Andrea, Elba, Italy	Host-associated	Annelida	Integument	Unclassified	MQ	Bacteria ; Proteobacteria	68.92	1.61	6131798	1	1	1	48	6484	602
☐	3300003129_3	Olavius algarvensis microbial community from off the coast of Capo di Sant Andrea, Elba, Italy	Host-associated	Annelida	Integument	Unclassified	MQ	Bacteria ; Proteobacteria ; Gammaproteobacter	91.99	2.73	4029764	0	0	0	34	4350	239

The table of bin scaffolds lists the scaffold IDs for the selected bin and links to table of genes for the selected bin scaffold.

Metagenome Bin Scaffolds

Genome: Olavius algarvensis microbial communities from Capo di Sant'Andrea bay, Elba, Italy - Delta1
Bin name: 2004178001_1

[Add Selected to Scaffold Cart] [Toggle Selected] [Select All] [Clear All]

Filter column: Scaffold ▼   Filter text ▼ : [   ] [Apply] ❓

[Export]   Page 1 of 2   << first < prev 1 2 next > last >>   100 ▼

[Column Selector]   [Select Page]   [Deselect Page]

Select	Scaffold ▲	Gene Count	Sequence Length (bp)	GC Content	Lineage	5s rRNA	16s rRNA	23s rRNA	tRNA Genes
☐	gws2_d1_0004_16	396	407380	0.49	Bacteria ; Proteobacteria ; Deltaproteobacteria ; unclassified ; unclassified ; unclassified ; Olavius algarvensis sulfate-reducing endosymbiont	0	0	0	1
☐	gws2_d1_0005_2	382	359424	0.49	Bacteria ; Proteobacteria	0	0	0	2
☐	gws2_d1_0006_16	270	312075	0.51	Bacteria ; Proteobacteria ; Deltaproteobacteria ; unclassified ; unclassified ; unclassified	0	0	0	1
☐	gws2_d1_0007_7	255	267432	0.49	Bacteria	0	0	0	1
☐	gws2_d1_0008_4	250	265094	0.50	Bacteria ; Proteobacteria ; Deltaproteobacteria ; unclassified ; unclassified ; unclassified ;	0	0	0	0

Metagenome Bin rRNA_5S Genes in Scaffolds

Genome: Olavius algarvensis microbial communities from Capo di Sant'Andrea bay, Elba, Italy - Delta1
Bin name: 2004178001_1

Filter column: Gene ID ▼   Filter text ▼ : [   ] [Apply] ❓

[Export]   Page 1 of 1   << first < prev 1 next > last >>   All ▼

[Column Selector]   [Select Page]   [Deselect Page]

Select	Gene ID ▲	Locus Type	Locus Tag	Gene Product Name	Start Coord	End Coord	Strand	Scaffold
☐	rRNA_1	rRNA	rRNA_1	5S rRNA	206776	206896	-	gws2_d1_0013_13

[Export]   Page 1 of 1   << first < prev 1 next > last >>   All ▼

[Add Selected to Gene Cart] [Select All] [Clear All]